Neuropsychology Review (2019) 29:498–521
https://doi.org/10.1007/s11065-019-09416-5

REVIEW

A Systematic Review of Neurocognitive Functioning in Behçet's Disease

Caroline A. Fisher[1,2] · Coco Bernard[1,3,4]

Received: 28 February 2018 / Accepted: 21 August 2019 / Published online: 7 September 2019

Abstract

Behçet's disease (BD) is a vascular, inflammatory multisystem disorder with neuro-Behçet's (NBD) diagnosed in a subset of patients with neurological manifestations. The objective of this review was to determine whether neurocognitive dysfunction is observed in BD, in which neurocognitive domains, and whether there are differences in rates of dysfunction observed between BD and NBD groups. Studies of any methodology were included that reported results from standardized neurocognitive assessment measures in participants with BD or NBD. Twelve group comparison studies met the criteria for inclusion in the review (totalling 284 BD and 157 NBD participants), as well as 17 case study/series papers (11 BD, 35 NBD). Issues with blinding, incomplete data reporting and selective reporting bias were found across the group and case study/series papers, as well as inadequate statistical adjustment for multiple comparisons in the group studies, and the lack of the use of appropriate norms or adjustment for premorbid ability in the case series/studies papers. These quality issues impacted on the conclusions that could be drawn from the current literature. Neurocognitive dysfunction was found in NBD compared to health controls (HC) in a higher proportion of results across studies, than in comparisons between BD and HC groups. The domains in which neurocognitive attenuation was most often reported were visual spatial ability, working memory and acquired knowledge, with more than 25% of these results showing significantly lower functioning in both the BD and NBD groups compared to HC. More than 25% of the processing speed and long-term memory encoding and retrieval results were also lower for the NBD group, compared to HC. Group comparisons between NBD and multiple sclerosis participants indicated few significant differences in neurocognitive test results. The majority of case study/series participants were found to have some degree of attenuated neurocognitive functioning, as defined by case study/series authors.

Keywords Behçet's · Cognition · Neuro-Behçet's · Neurocognitive · Neuropsychology

Behçet's disease (BD) is a vasculitis and multisystem inflammatory syndrome (Kalra et al., 2014; Dalvi, Yildirim, & Yazici, 2012). While there is no definitive single medical investigation that can be used to confirm the disease (Kronborg, Mahar, & Kelly, 2014), defined clinical criteria exist to aid diagnosis (e.g., Davatchi et al., 2014; International Study Group for Behçet's Disease, 1990). The most characteristic features of BD are oral ulcerations, genital ulcerations and skin lesions, followed by eye lesions and abnormalities in the vascular and central nervous systems (Davatchi et al., 2011). BD has a variable regional prevalence, being most common in Middle Eastern and eastern Mediterranean countries followed by Asia (Hatemi, Yazici, & Yazici, 2013). The aetiology remains unclear. A genetic component has been suggested as well as potential environmental or biological triggers (Dalvi et al., 2012; Ideguchi et al., 2010; Remmers et al., 2010). A variety of pharmacotherapies are presently used to treat the symptoms of the disease, including corticosteroids, anti-inflammatory agents, mercaptopurine derivates, calcineurin inhibitors, and tumour necrosis factor (TNF)-blocking agents, and treatment guidelines are now available

✉ Caroline A. Fisher
Caroline.Fisher2@mh.org.au

[1] Allied Health - Psychology, Melbourne Health, Royal Melbourne Hospital, 300 Grattan St, Parkville, Melbourne, Victoria 3052, Australia

[2] Neuropsychology Service, The Melbourne Clinic, Healthscope, Richmond, Melbourne, Australia

[3] Adult Neuropsychology Service, Rehabilitation and Aged Care Services, Monash Health, Kingston Centre, Cheltenham, Australia

[4] Adult Neuropsychology Service, Rehabilitation and Aged Care, Monash Health, Caulfield Hospital, Caulfield, Australia

(Alibaz-Oner, Sawalha, & Direskeneli, 2018; Kikuchi et al., 2017; Ozguler & Hatemi, 2016; Türsen & Türsen, 2014).

Central nervous system (CNS) involvement occurs in BD, although rates vary between 2.3 and 44% of cases across studies (Davatchi, Shahram, Chams-davatchi, Shams, Nadji, Akhlaghi, et al., 2010; Davatchi, 2012; Tursen, Gurler, & Boyvat, 2006; Bang et al., 2001; Madanat, Fayyad, Verity, & Zureikat, 2000; Valesini, Pezzi, Catarinelli, Accorinti, & Priori, 1991; Al-Dalaan et al., 1994; Domingos et al., 2015). Differences in the criteria and screening/assessment methods used to detect neurological and CNS abnormalities, as well as regional variations are likely to contribute to this variation. Patients with neurological involvement are categorised with the sub-syndrome of neuro-Behçet's disease (NBD). Once a recognised set of criteria for BD are met, NBD is diagnosed if there is are co-occurring neurological abnormalities believed to be caused by the BD (Kalra et al., 2014; Siva & Saip, 2009). This can include a range of sequelae such as parenchymal features (brainstem or multifocal lesions, myelopathy, cerebral abnormalities, or optic neuropathy) and non-parenchymal features (cerebral venous thrombosis, intracranial hypertension syndrome, or acute meningeal syndrome), with parenchymal presentations more common. Headache is also often reported in BD and has been found to occur with or without documented neurological abnormalities (Farahangiz, Sarhadi, Safari, & Borhani-Haghighi, 2012; Kale, Agaoglu, Icen, Yazici, & Tanik, 2008). Differing syndromes within the NBD cohort have been previously separated into three categories: patients with increased intracranial pressure (with or without cerebral venous sinus thrombosis), patients presenting with a stroke, and patients presenting with spinal cord involvement and CSF pleocytosis. Brainstem involvement has been noted to be independent of group (Shakir, Sulaiman, Kahn, & Rudwan, 1990). NBD can be monophasic, polyphasic or progressive in its course (Farahangiz et al., 2012; Haghighi, Sarhadi, & Farahangiz, 2011).

Neuroimaging studies have indicated that the most frequently affected brain regions in NBD are the superficial and periventricular white matter, the midbrain, and pons (Farahangiz et al., 2012), while in non-parenchymal NBD cerebral venous sinus thrombosis is most commonly observed (Farahangiz et al., 2012). NBD has also been shown to significantly impact on quality of life, with a longitudinal study finding that 46% of participants with NBD had at least a moderate level of functional disability based on the modified Rankin Scale (Noel et al., 2014) with neurological, physical, and psychological sequelae contributing to the experienced disability (Borson, 1982; Matsui et al., 2010; Noel et al., 2014; Özdemir, Özsoylar, Candansayar, Coşar, & Önder, 2004; Uğuz, Dursun, Kaya, & Cilli, 2007).

Neurocognitive functions, including, attention, information processing, memory, language, visual-perceptual processing, reasoning, impulse control, planning and organisation are a group of skills that are localised within the brain, and can be affected by acquired neurological damage. Several recent narrative reviews have suggested that neurocognitive deficits in BD/NBD primarily occur in the areas of attention, working memory, delayed recall/retrieval, and executive functioning (Fisher, Sewell, & Baker, 2016; Kidd, 2017). Past research has also revealed that attenuated functioning can be seen in the absence of neurological manifestations (Erberk-Ozen, Birol, Boratav, & Kocak, 2006; Monastero et al., 2004; Zayed et al., 2011). However, a thorough review of the total literature has yet to be conducted, and it remains unclear if there are differences in the neurocognitive profiles of BD compared to NBD patients. The purpose of this review was to systematically examine the literature to determine what is known about neurocognitive dysfunction in BD and NBD. This included examining the rates, nature, severity and course of neurocognitive impairment in people with BD and NBD. Given the unknown size of the study pool prior to conducting the search, studies of all methodologies were included (i.e., group comparisons, group characterisations, case series, and case studies), provided they reported on the results of standardized measures of neurocognitive functioning.

Method

The systematic review was conducted in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Liberati et al., 2009; Moher, Liberati, Tetzlaff, & Altman, 2009) and adapted for neuropsychological research (Gates & March, 2016). A review protocol was registered on the International Prospective Register of Systematic Reviews (PROSPERO ID: CRD42017072840) on the 24th of July 2017. Included in the review was any study that investigated neurocognitive functioning in individuals diagnosed with BD or NBD. Neuropsychological investigation was defined as any standardized measure of neurocognitive functioning, including brief screening assessments, specific neurocognitive tests, and large neuropsychological assessment batteries. Standardised self and informant report measures of cognitive functioning were also included. Studies with any methodology were included, encompassing group comparison studies with healthy controls, group studies comparing two or more clinical groups, group categorization studies, longitudinal cohort follow-up studies, case series and case studies.

All of the following criteria were required for inclusion in the review:

1) Original data pertaining to human patient(s) with Behçet's disease (using any criteria for diagnosis)
2) Any study methodology reporting empirical data (originating in or based on observation or experience)

3) Reporting of results on standardized measures of neurocognitive/neuropsychological assessment in either data format (raw scores, group means and standard deviations), scaled scores format (e.g., ranging from 1 to 19), percentiles that pertain to normative data, score ranges appropriate to the neurocognitive assessment being used (e.g., Extremely Low, Low, Borderline, Low Average, Average, High Average, Superior, Very Superior), or via the reporting of differences in mean scores between a group of Behçet's participants and a control or comparison group (, p values of significance)

Studies were excluded if they met any of the following criteria:

1) The data of the patient(s) with Behçet's disease could not be separated from the data of patients with other diseases/disorders
2) Neurocognitive/neuropsychological assessment measures were described as having been administered, but the results were not reported
3) Neurocognitive/neuropsychological assessment measures were described as having been administered, and were described in vague or generalized terms that could not be interpreted in reference to any form of raw scores, ranges, standard scores or normative data (e.g., "no impairment was found,, or "patient was impaired").

The PsycINFO, Embase/Medline and PubMed databases were all searched (April 18th 2017), with the terms Behcet (Behcet*) combined with cognitive, cognition, neuropsychological, neuropsychology or neurocognitive. There were no restrictions on language. All identified studies were independently screened by two doctoral level clinical neuropsychologists. The independent screening ratings were then collated and revealed a 94.2% inclusion/exclusion agreement rate. All discrepancies were resolved by discussion between the two study authors, until mutual agreement was reached. The relevant data was then extracted from all included studies. Data was extracted by one study author and the accuracy of all extracted data was independently cross-checked by the second author. Risk of bias and quality ratings were independently conducted by both authors, across all included papers. Ratings were then collated and cross-checked with discrepancies resolved via discussion until mutual agreement was reached.

A risk of bias analysis was undertaken using a modified framework based on the Cochrane Handbook for Systematic Reviews of Interventions (Higgins, Altman, & Sterne 2017). Bias ratings pertaining solely to interventions were removed from the standard Cochrane framework (i.e. selection bias and performance bias), with only bias ratings relevant to the reporting of characterisation studies included (i.e. detection bias, attrition bias and reporting bias). As an additional quality measure, raters also indicated whether authors had utilised appropriate adjustments for multiple statistical comparisons (i.e. to reduce the Type I error rate). For the case series/studies two additional quality measures were included; 1) appropriate use of, or reference to, normative data, 2) whether authors considered participant's premorbid ability level when determining the nature/level of impairment. Where the quality of reporting differed for cases reported within a paper, these cases were separated so that quality and bias ratings were provided for each case within the series.

Results

Study Selection

Figure 1 outlines the results from the search, as defined by aforementioned search criteria. All studies were full text screened, except in instances where only abstracts were available ($N = 14$) and foreign language articles. Of the foreign language articles, 11 abstracts and five full text articles were translated. Attempts were made to contact authors to obtain missing data when cognitive assessments/data were referred to, but not reported, or when studies did not report all of their cognitive assessment results. This did not yield any additional useable data.

A total of 29 studies met the inclusion criteria and were included in the review with study design breakdowns noted in Fig. 1. Group comparison studies were comprised of nine studies which compared BD/NBD to a healthy control group and three studies comparing BD/NBD participants with multiple sclerosis (MS). One study was presented as a group comparison (NBD vs MS) and as a NBD group categorisation, across two separate papers (Bingol et al., 2011; Topcular et al., 2011). As such, results are presented across both the group comparison and group categorisation sections. Overall, the included group studies contained a total of 284 BD participants, 157 NBD participants, 297 healthy control participants and 119 participants with MS. The included case study/series papers contained a total of 46 cases, 11 with BD and 35 with NBD.

Group Comparison Studies

Study Characteristics, Quality and Bias Ratings

Demographic features and study details of the 12 included group comparison studies (reported across 13 papers) are outlined in Table 1. This geographical distribution roughly aligns with known high prevalence areas of BD, except for the lack of studies from eastern Asia. Sample sizes for the studies were generally small, ranging from 10 to 78. Mean

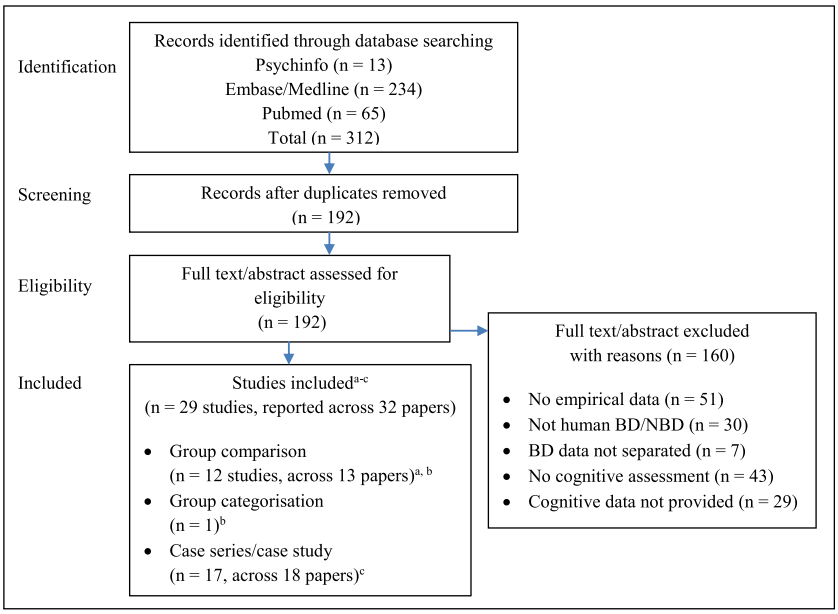

Fig. 1 Flow of papers through phases of the systematic review by phase

[a] One study was published twice (BD vs Healthy Controls), with marginally different statistical analysis presented in each.

[b] One study was presented across two papers; 1) a group comparison (NBD vs MS); 2) isolated characterisation of the NBD group, with increased detail regarding NBD neurocognitive test results.

[c] One study was published twice (in English and Japanese), with relevant data extracted from both papers.

ages for all groups hovered around the 30's and 40's. Gender ratios varied across studies and comparison groups, with four studies reporting significant gender imbalance across comparison groups. The disease duration of the BD/NBD participants varied between 6.15 and 14.9 years (where reported). Eight studies reported the medication status of their BD/NBD participants. Only one study (Nurova, Kurtuncu, Coban, Birday, & Eraksoy, 2014) described longitudinal assessment follow-up post a cognitive rehabilitation program, although not all of this data was reported. The majority of studies (11) reported having conducted some form of psychological assessment, in addition to the neurocognitive assessment. The most commonly assessed psychological conditions were depression and anxiety.

The majority of group comparison studies (10 of 12) utilised the International Study Group for Behçet's disease (1990) criteria, although some variations were accepted in one study (i.e. Hernandez, 2002 indicated that two of their participants who did not meet criteria satisfied clinical conventions for diagnosis utilising other diagnostic frameworks; Mason & Barnes, 1969 and O'Duffy & Goldstien, 1979). The two remaining studies did not state which diagnostic criteria they used. Further diagnostic distinction was required for those studies that included both NBD and BD participants. In three of the studies. Authors provided information about their methods accordingly (Cavaco et al., 2009; Dutra et al., 2013; Sucullu Karadag et al., 2014). In a fourth study, details about the diagnosis of NBD was provided but it was unclear if NBD was explicitly ruled out in the BD participants (Erkol, Vural, Karantay, & Uluduz, 2009). A further study comparing BD to HC provided information about how NBD was excluded in the BD participants (Monastero et al., 2004). NBD diagnostic criteria were also provided in the Bingol et al. (2011) study comparing NBD to MS. In the two remaining studies that included NBD participants, explicit methodology for diagnosing NBD was not reported, however both stated that the NBD cases all had parenchymal involvement. Six studies explicitly indicated exclusion criteria for their BD/NBD participants.

Table 2 displays the risk of bias and quality summary ratings for the group comparison studies. The blinding status of the neurocognitive assessor was unclear in two-thirds of the studies. Detection bias for 'incomplete data reporting' was considered low risk in two-thirds of studies, was unclear in a quarter of studies, and deemed high risk' in one study. The majority (11 of 12) studies rated low on criteria for selective reporting bias. Appropriate adjustments to p values for multiple statistical comparisons had only clearly occurred in a quarter of studies ($n = 3$). In the remaining studies, six were deemed unclear as to whether adjustments had been made and three provided no evidence of appropriate adjustment to p values. Classifications for 'other bias ratings' were high in eight studies. Identified issues included inaccurate interpretation of the study results, findings that were reported differently throughout the manuscript, demographic differences in the comparison populations, differing group numbers reported throughout the manuscript, and/or if neurocognitive test versions were reported incorrectly. Only one paper was rated as

Table 1 Group comparison studies included in systematic review

First Author, Year	Country	Partici-pants	Mean age (SD)	F:M	Mean disease dur-ation (SD)	BD Exclusion criteria provided	BD medication status	Psychological assessment measures
Nurova et al., 2014	Turkey	33 NBD 59 MS	–	–		–	–	BDI
Sucullu Karadag et al., 2014	Turkey	34 BD 11 NBD 32 HC	36.8 (9.3) 39.1 (10.4) 34.4 (9.7)	44:56 27:73 50:50	7.5 (7.2) 11.5 (6.5) -	BDI ≥17, missing clinical data, difficulty cooperating on tests	87% of patients on colchicine, 63.3% had used corticosteroids with 22 were still using a dose of 12 ± 12.8 mg/day	BDI administered as part of screening, but scores not reported
Dutra et al., 2013	Brazil	24 BD 24 NBD 24 HC	39.4 (10.8) 39.5 (10.5) 41.9 (9.2)	38:62 61:39	10.37 (8.95) 6.27 (4.96) -	Amaurosis, < 3 years formal education, contraindications for brain MRI	Cumulative prednisolone was reported for BD and NBD patients. BD 13.74 g (15.22), NBD 12.38 (10.8) gm	HAM-A, BDI, Epworth Sleepiness Scale
Gunduz, 2012	Turkey	20 NBD 20 MS	39.3 (10.5) 36.7 (9.5)	35:65 75:25	6.2 (5.1) 7.6 (5.4)	Severe visual loss, illiteracy, physical/ cognitive disability hampering cognitive testing, < 2 months after most recent attack.	Patients on intravenous steroid treatment excluded, but no details regarding medication status of those included was provided.	BDI, NPI
Zayed et al., 2011	Egypt	25 BD 10 HC	33.8 (8.7) 31.9 (n/a)	24:76 30:70	7 (n/a)	–	22 oral prednisolone, 8 colchicine, 8 IV cyclophosphamide pulses; 11 azathioprine, 10 oral anti-coagulants, 2 meth-nonsteroidal anti-inflammatory drugs, 2 methotrexate, 1 chlorambucil	HARS, HDRS
Bingol et al., 2011	Turkey	40 NBD 40 MS	38.1 (10.2) 37.2 (9.8)	40:60 67:33	14.9 (8) 11.9 (9.1)	Disease activity or pulse steroids within the last 30 days, concomitant psychiatric disorder. < 18 or > 65 years, less than primary school education.	–	–
Shahram et al., 2010	Iran	65 BD 43 HC	- -	- -	- -	–	Prednisolone use for some.	BAI, HDRS
Cavaco et al., 2009	Portugal	15 NBD 32 HC1 35 BD 78 HC2	48.2 (10.5) 47.4 (10.1) 40.7 (9.9) 40.55 (11.4)	67:33 69:31 77:23 68:32	14.4 (7.73) - 13.91 (8.28) -	–	NBD: 4 Prednisolone, 1 Anti-anxiety medication, 3 antidepressants BD: 14 Prednisolone, 1 Anti-anxiety medication, 1 antidepressants	HADS (for BD and NBD only)
Erkol et al., 2009	Turkey	BD 14 NBD 14 HC 12	35.9 (8.2) 36.7 (8.3) 32.2 (5.3)	43:57 36:64 34:66	7.1 (5.3) 7.0 (8.6)	Psychiatric disorder, other systemic disease (e.g. diabetes, chronic renal failure), the use of any drug in last 1 week.	–	HDRS - no findings reported
Erberk-Ozen et al., 2006 & Ozen 2004	Turkey	30 BD 30 HC	34.5 (9.1) 33.9 (8.7)	27:63 57:43	8.9 (7.3) -	<20 or > 55 yrs., less than graduated primary school education, medical/mental handicaps that could influence test performance negatively,	14 Colchicin, 7 Nonsteroidal treatment + colchicine, none were on systemic corticosteroids	DMS-IV Assessment, BDI

Table 1 (continued)

First Author, Year	Country	Partici-pants	Mean age (SD)	F:M	Mean disease dur-ation (SD)	BD Exclusion criteria provided	BD medication status	Psychological assessment measures
						neurological or other disease, history of alcohol or substance abuse.		
Monastero et al., 2004	Italy	26 BD 26 HC	33.2 (9.6) 32.8 (9.7)	50:50 50:50	6.15 (5.4) -	–	2 Ciclophosphamide, 2 Non-steroidal anti-inflammatory drugs, 7 Cyclosporin A, 13 Colchicine, 14 Prednisone	HARS, HDRS
Hernandez, 2002	Spain	31 BD 10 HC	–	–	–	–	–	MMPI

Abbreviations: BD Behcet's disease, *BDI* Beck Depression Inventory, *DSM-IV* Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition), *HARS/HAM-A* Hamilton Anxiety Rating Scale, *HC* Healthy Controls, *HDRS/HAM-D* Hamilton Depression Rating Scale, *MRI* Magnetic Resonance Imaging, *MMPI* Minnesota Multiphasic Personality Inventory, *NPI* Narcissistic Personality Inventory, *NBD* Neuro-Behcet's disease

having a low risk of bias across all areas assessed, as well as appropriate adjustment for multiple comparisons, and even within this paper a typographical error was found within the results section stating HARS and HDRS results went in the opposite direction to what reported in the papers tables.

Neurocognitive and Psychological Assessment Data

Table 3 shows the neurocognitive assessment results of the group comparison studies. There was a high degree of variation in the size of the cognitive testing batteries applied, from fairly extensive assessment batteries, examining multiple domains (e.g., Cavaco et al., 2009; Dutra et al., 2013; Gündüz et al., 2012; Monastero et al., 2004; Sucullu Karadag et al., 2014) to studies that employed just one neurocognitive assessment measure (e.g.,e.g.,, Erkol et al., 2009). Seventy five percent of articles described the complete set of neurocognitive results.

Table 4 presents the neurocognitive assessment outcomes from the comparisons group, significance levels, and effects sizes. Effects sizes were calculated via the standard Cohen's *d* and Hedges' *g* calculations (Cohen, 1988; Larrabee, 2014). Where multiple results were reported for a measure, total or overarching scale scores were included, rather than individual sub-test scores. This was to ensure the most robust measures were represented and to avoid duplication of results. For example, if the Memory Quotient (MQ) scores from the Wechsler Memory Scales (WMS) were available then these were used instead of individual subtests scores.

There is presently no universally accepted method for grouping or separating neurocognitive test results by domain or subdomain. This is evidenced by the differing neurocognitive domain divisions utilised in recent systematic reviews examining neurocognitive functioning in particular disorders (Beeldman et al., 2016; Evans, Iverson, Yatham, & Lam, 2014; Ganzer, Bröning, Kraft, Sack, & Thomasius, 2016; Goodall et al., 2018). Rather, the grouping of neurocognitive test results and allocation of specific neurocognitive tests to specified domains is generally conducted according to established practice and clinical convention. It has been recently suggested by Jewsbury and colleagues, however, that classifying tests based on construct validity (determined by the large body of psychometric research into intelligence and cognitive ability) may be both more valid and useful (Jewsbury, Bowden, & Duff, 2017). They indicated that the Cattell-Horn-Carroll (CHC) model of cognitive individual difference can be applied to clinical populations as the same cognitive constructs are reflected in test scores in clinical populations as in community and educational samples. In line with this research, the neurocognitive test results were allocated to one of the broad CHC constructs according to the recently described classification system (Jewsbury et al., 2017; Jewsbury, Bowden, & Strauss, 2016;

Table 2 Risk of bias and quality rating summary table for group characterisations studies

First Author, Year	Blinding of Cognitive Assessor (Detection Bias)	Incomplete Data Reporting (Detection Bias)	Selective Reporting Rating (Reporting Bias)	Other Bias Rating	Appropriate adjustment for multiple statistical comparisons	Comments
Nurova et al., 2014	Unclear	Unclear	High	Unclear	Unclear Only p values (not results reported), however an adjustment does not appear to have been applied.	Actual outcome data is not provided, only p values and discussion of significance levels. Comments made in abstract about inhibition results are not supported by information results section.
Sucullu Karadag et al., 2014	Unclear	Unclear	Low	High	Unclear Both corrected and uncorrected results reported.	Statements in discussion that are not supported by the results after Bonferroni adjustment. Inconsistencies in inclusion numbers between abstract and main text.
Dutra et al., 2013	Low	Low	Low	High	Unclear A Bonferroni correction is reported, but does not appear to have been applied correctly.	Bonferroni correction does not have sufficient *p*-value adjustment for the 21 comparisons. Inconsistency in test versions reported throughout text and references.
Gunduz, 2012	Low	Low	Low	Unclear	Unclear Both corrected and uncorrected results reported.	Tables and findings reported were in reference to unadjusted p-value cut-off of 0.05, so findings were misleading. Gender bias between groups.
Zayed et al., 2011	Unclear	Low	Low	High	No	No adjustment for multiple comparisons appears to have been made to results. As a result a number of conclusions drawn from the results appear to be incorrect.
Bingol et al., 2011	Unclear	Low	Low	High	Unclear Not stated, but significant *p*-values all at levels consistent with appropriate correction.	Large gender imbalance in comparison groups. Authors have drawn conclusions about whole cognitive domains from single tests (e.g.

Table 2 (continued)

First Author, Year	Blinding of Cognitive Assessor (Detection Bias)	Incomplete Data Reporting (Detection Bias)	Selective Reporting Rating (Reporting Bias)	Other Bias Rating	Appropriate adjustment for multiple statistical comparisons	Comments
						word list generation inferred to reflect 'executive functions' as a whole)
Shahram et al., 2010	Unclear	Low	High	Unclear	Unclear Not reported, but no significant differences found on neurocognitive measures	Participant characteristics and not all test results fully reported. Conclusions about whole cognitive domains from single tests (e.g. WCST used as measure of 'visuospatial ability', and block design inferred to reflect 'executive functions' as a whole).
Cavaco et al., 2009	Unclear	Unclear	Low	High	No	No adjustment for multiple comparisons appears to have been made to results. As a result a number of conclusions drawn from the results appear to be incorrect.
Erkol et al., 2009	Unclear	Low	High	High	Yes	Information reported in the abstract states the opposite of the results reported in the results section of the paper, i.e. "BS group displayed similar electrophysiological changes and PASAT results compared to controls".
Erberk-Ozen et al., 2006; Ozen 2004	Unclear	Low	Low	High	No (Ozen 2004) Unclear (Erberk-Ozen et al., 2006) - no information provided about adjustments conducted as part of ANCOVA, however an adjustment does not appear to have been applied.	No adjustment for multiple comparisons appears to have been made to results. As a result a number of conclusions drawn from the results appear to be incorrect. Misleading descriptions of results in other areas also.
Monastero et al., 2004	Low	Low	Low	Low	Yes	There is an error in the paper in the results

Table 2 (continued)

First Author, Year	Blinding of Cognitive Assessor (Detection Bias)	Incomplete Data Reporting (Detection Bias)	Selective Reporting Rating (Reporting Bias)	Other Bias Rating	Appropriate adjustment for multiple statistical comparisons	Comments
						section. States: "Patients with BD showed significantly lower scores than controls on the HARS and on the HDRS. In fact, the scores are higher, not lower.
Hernandez, 2002	Low	High	High	High	Unclear No significant differences reported	Participant numbers change throughout the paper. Raw test scores not reported, 25% missing data reported. WAIS referred to as a memory test.

Abbreviations: *ANCOVA* Analysis of covariance, *HARS* Hamilton Anxiety Rating Scale, *HDRS* Hamilton Depression Rating Scale, *PASAT* Paced Auditory Serial Addition Test, *WCST* Wisconsin Card Sorting Task, *WAIS* Wechsler Adult Intelligence Scale

Hoelzle, 2008; Jewsbury & Bowden, 2017). In cases where the neurocognitive test has shown to load onto more than one factor (e.g., Clock Drawing Test, Logical Memory, Matrix Reasoning), the construct for which the test has most consistently the strongest loading was chosen. For transparency, Table 4 provides an overview of the CHC constructs and all tests subsumed within each construct. None of the included studies utilised neurocognitive tests that corresponded to the Gq (Quantitative ability) or Ga (Auditory ability). In keeping with the factor structure research, 'executive functioning' tasks involving switching, inhibition and updating, were not classified as a separate construct, as there is little evidence to support this within the CHC model. Rather, updating tasks were grouped with general memory factors (Glr), and inhibition and switching tasks were grouped with general speed (Gs) (Jewsbury et al., 2016). All traditionally classified 'executive functioning' tests have been found to load more strongly on to other existing CHC constructs, and we have therefore classified these tasks as such. Tests based solely on subjective ratings of cognitive functioning (e.g., VAS, FBI) were not included in this section of the analysis.

Eight studies contributed results to the BD versus HC comparison, totalling 258 BD participants and 229 HC participants (Cavaco et al., 2009; Dutra et al., 2013; Erberk-Ozen et al., 2006; Erkol et al., 2009; Hernández et al., 2002; Shahram et al., 2010; Sucullu Karadag et al., 2014; Zayed et al., 2011). Across all domains, 82% of reported results suggested non-significant differences between groups. The most significant group differences (reported in more than 25% of results) were in the domains of acquired knowledge/crystallised ability, working memory and visual spatial ability. On measures of psychological functioning, the four studies that investigated depression found higher rates of depression in BD compared with HC (Dutra et al., 2013; Erberk-Ozen et al., 2006; Monastero et al., 2004; Zayed et al., 2011). Similarly, rates of anxiety were significantly higher in BD compared to HC participants in the three studies that included anxiety measures. A further study controlled for depression and anxiety in their analysis of neurocognitive outcomes but did not report on these measures directly (Shahram et al., 2010).

Four studies compared NBD versus HC groups, totalling 64 NBD participants and 100 HC participants. (Cavaco et al., 2009; Dutra et al., 2013; Erkol et al., 2009; Sucullu Karadag et al., 2014). Across all domains, 69% of results indicated non-significant group differences. The most significant group differences (>25% of results) were across the domains of acquired knowledge/crystallised ability, processing speed, long term memory encoding and retrieval, working memory and visual spatial ability. The one study that assessed psychological symptoms found significantly higher rates of both depression and anxiety in NBD compared to HC (Dutra et al., 2013).

A total of four studies compared NBD with BD participants (Cavaco et al., 2009; Dutra et al., 2013; Cavaco et al., 2009;

Table 3 Neurocognitive assessment results of group comparison studies

First Author, Year	Measures with no significant differences	Measures with significant differences	Measures described but results not reported
Nurova et al., 2014 [a]	NBD vs MS: Digit span test	None	ACE, 10/36 spatial recall test, SRT, Burdon attentive
Sucullu Karadag et al., 2014	NBD vs HC: Self-report VAS (cognition), Stroop – interference, SDMT, FST BD vs HC: Self-report VAS (cognition), Stroop – interference, Stroop – spontaneous correction, SDMT PASAT, JLOT, FST, TMT-A, TMT-B NBD vs BD: Self-report VAS (cognition), RAVLT – immediate memory, RAVLT – delayed free recall, ACTT, Stroop – interference, COWAT, SDMT, PASAT, JLOT, FST, TMT-A, TMT-B, DST	NBD worse than HC: RAVLT – immediate memory, RAVLT – learning, RAVLT – delayed free recall ACTT, Stroop - spontaneous corrections, Stroop – errors, COWAT, PASAT, JLOT, TMT-A, TMT-B, DST BD worse than HC: RAVLT - immediate memory, RAVLT – learning, ACTT, Stroop – errors, COWAT, DST NBD worse than BD: RAVLT – learning, Stroop – spontaneous corrections, Stroop – errors	None
Dutra et al., 2013	BD vs NBD vs HC: MMSE, DS – forwards, DS – backwards, VPA – first trial, VPA – second, VPA – third trial, VPA – delayed trial, COWAT – phonemic or semantic, TMT-A, TMT-B, TMT B-A, WAIS Matrix Reasoning BD vs HC: ROCF immediate recall, ROCF delayed recall NBD vs BD: CDT, BNT	BD worse than HC: ROCF Copy, BNT (15-item), CDT NBD worse than HC: ROCF immediate recall, ROCF delayed recall, BNT (15-item), CDT NBD worse than BD: ROCF immediate recall, ROCF delayed recall	None
Gunduz, 2012 [a]	NBD vs MS: PASAT, Nine-Hole Peg Test (Left or Right), WAIS-R DST – Forward, WAIS-R DST – Backwards, CPT – Hits, CPT – omissions, CPT – comissions, CVLT – trials 1–4, CVLT –trial 5, CVLT – short delayed free recall, CVLT – long delayed free recall, CVLT – short delayed cued recall, CVLT – long delayed cued recall, CVLT – semantic clustering, CVLT – perseverations, CVLT – recall intrusions, CVLT – recognition hits, Stroop – corrected interference time, Stroop – errors, Stroop – spontaneous corrections, COWAT – animals, COWAT (letters 'K', 'A', 'S'), WCST – number of trials to complete first category, WCST – number of completed categories, WCST – number of perseverations, WCST – failure to maintain set score, WCST – conceptual level response percentage, HVOT, Facial Recognition, JLOT, BNT - correct responses, BNT – total (correct + cued responses), FBI – total score, FBI – negative score, FBI – disinhibition	None	None
Zayed et al., 2011 [a]	BD vs HC: WMS-R[b] – Orientation, Mental Control, Memory Quotient, Paired Associate Language, Logical Memory, Information	BD worse than HC: WMS-R[b] - Digit Span, Visual Reproduction	None
Bingol et al., 2011	NBD vs MS: SDMT	NBD worse than MS: PASAT, SRT, SPART, WLG	None
Shahram et al., 2010	BD vs HC: WMS, WAIS - Block Design, WCST	None	None
Cavaco et al., 2009 [a]	NBD vs HC1: 9-hole peg test, Attentive matrices test, WAIS DST – forwards, WAIS DST – backwards, Corsi-block tapping test, Sentence repetition, RCFT-copy, RCFT-30 min recall, RAVLT – immediate recall, RAVLT-30 min recall, RAVLT-30 min recognition, COWAT (letters 'C', 'A', 'E'), WCST – number of categories, WCST – number of perseverative errors, TMT-A, TMT-B BD vs HC2: 9-hole peg test, Attentive Matrices Test, Corsi-block Tapping Test, Sentence Repetition, RCFT – copy, RAVLT – immediate recall, RAVLT – 30 min recall, RAVLT – 30 min recognition, WCST – number of categories, WCST – number of perseverative errors, TMT-A, TMT-B, WAIS DST – backwards, COWAT	BD worse than HC2: WAIS DST – forwards	None
Erkol et al., 2009	None	BD worse than HC: PASAT NBD worse than HC: PASAT	MMSE used at screening
Erberk-Ozen et al., 2006 [a] and Ozen 2004	BD vs HC: WCST – total errors (%),WCST – number of perseverative errors (%), WCST – conceptual level responses (%), WCST – failures to maintaining set, Stroop – part 1, Stroop – part 2, Stroop – part 3, Stroop – part 4, Stroop – part 5	BD worse than HC: WCST – categories completed	None

Table 3 (continued)

First Author, Year	Measures with no significant differences	Measures with significant differences	Measures described but results not reported
Monastero et al., 2004	BD vs HC: WAIS Digit Span, Corsi-Block Tapping Span, RAVLT - delayed recall (10 min), RCFT – copy, Token Test, TMT – A, TMT – B, Raven's Coloured Progressive Matrices, Phonemic Fluency, Category Fluency	BD worse than HC: RAVLT - Immediate Recall, RCFT – 10 min recall, JLOT	MMSE used at screening
Hernandez, 2002	BD vs HC: Square test of Letters, WAIS, Thurstone's PMA	None	None

Abbreviations: *ACE* Addenbrook's Cognitive Examination, *ACTT* Auditory Consonant Trigram Test, *BD* Behcet's disease, *BNT* Boston Naming Test, *CDT* Clock Drawing Task, *CPT* Continuous Performance Test, *COWAT* Controlled Oral Word Association Test, *CVLT* Californian Verbal Learning Test, *DST* Digit Span Test, *FBI* Frontal Behavioural Inventory, *FST* Faces Symbol Test, *HC* Healthy Control, *HVOT* Hooper Visual Organisation Test, *JLOT* Judgement of Line Orientation Test, *SPART* Spatial Recall Test, *SRT* Serial Reminding Test, *JLOT* Judgement of Line Orientation Test, *MS* Multiple Sclerosis, *NBD* Neuro-Behcet's disease, *PASAT* Paced Auditory Serial Addition Test, *PMA* Primary Mental Abilities Test, *RAVLT* Rey Auditory Verbal Learning Test, *ROCF/RCFT/ROCFT* Rey-Osterrieth Complex Figure Test, *SDMT* Symbol Digit Modalities Test, *TMT-A* Trail Making Test A, *TMT-B* Trail Making Test B, *VAS* Visual Analogue Scale, *WAIS* Wechsler Adult Intelligence Scale, *WCST* Wisconsin Card Sorting Task, *WLG* Word List Generation, *WMS* Wechsler Memory Scales

[a] Studies with inadequate adjustment for multiple comparisons, results only reported if p-values remain significant after reviewer applied Bonferroni correction

[b] Subtests reported do not correspond to WMS-R. These are actually from the WMS. Test reporting appears to be incorrect and is not referenced in paper

Sucullu Karadag et al., 2014); however, direct comparison of the NBD and BD was only possible in two studies, due to the study methodologies and statistical comparison analysis applied (Dutra et al., 2013; Sucullu Karadag et al., 2014). In these studies, 88 % of group comparison results were non-significant. These non-significant findings were prevalent across all construct domains except for visual-spatial ability where 50% of results indicated poorer visuo-spatial ability in NBD when compared with BD participants. One study reported on psychological assessment outcomes, with no significant differences found between NBD and BD participants on measures of anxiety or depression (Dutra et al., 2013).

The final group comparisons were between NBD and MS participants, with three contributing studies containing a total of 93 NBD participants and 119 MS participants (Bingol et al., 2011; Gündüz et al., 2012; Nurova et al., 2014). Eighty-nine percent of results across domains were consistent with non-significant group differences, and the majority of results were non-significant in all construct domain groups. The construct with the highest proportion of results with significant differences (NBD worse than MS) was FW Word fluency (33%). The one study that reported on psychological assessment results found no significant group differences in depression and neuropsychiatric symptoms between NBD and MS participants (Gündüz et al., 2012).

In addition to reporting group average scores, a number of studies categorised participants according to overall impairment rates or domain specific impairment rates, as compared to the control group (Cavaco et al., 2009; Dutra et al., 2013; Monastero et al., 2004). However, the criteria used to determine what constituted 'impairment' was different in each study. Dutra et al. (2013) placed their cut off at scores less than the 5th percentile of HC performance. Monastero et al. (2004) also utilised a 5th percentile cut-off but based this on performance that was less than the 5th percentile of the 'normal population' with raw scores corrected by age, education and gender. Cavaco et al. (2009) set the criteria for impairment as performances that were comparable to the weakest 1 % of the HC participants. As such, few comparisons could be drawn from the data presented in this manner.

Group Characterisation Studies

One study provided a characterisation of 40 NBD participants, reporting group means and standard deviations from a battery of neurocognitive tests comprised of PASAT, SRT, SPART, SDMT, and Word List Generation (Topicular, 2011; also reported in Bingol et al., 2011). Unfortunately, normative data for these tests were not provided and the degree of attenuation from normal or Average range functioning on these tasks was unclear. The authors state that the most commonly affected neurocognitive functions were sustained attention and information processing speed (58% of participants), visual-spatial

Table 4 Collated neurocognitive assessment outcomes from group comparison studies by comparison group and CHC constructs

Comparison Groups	Comparison outcomes by significance level and effect size[a]	G General Cognitive Ability: IQ (WAIS), MMSE	Gc Acquired knowledge/ crystallised ability: BNT, Token test, Thurstone's PMA	Gs Processing Speed: TMT A&B, Digit Symbol tests, Stroop, Coding, PASAT, FST, 9-hole peg test, Attentive Matrices	Glr Long-term memory encoding and retrieval: RAVLT, VPA, WMS MQ, CVLT, Facial Recognition, SRT, SPART	Gsm Working memory: DS (forwards, backwards, combined) Sentence repetition, Corsi Block Span, Square Test of Letters, ACTT, CPT	Gv Visual-spatial ability: RCFT/ROCF Copy & Recall, JLOT, WAIS BD, CDT, HVOT	Gf Fluid reasoning: WCST, WAIS MR, Ravens Progressive Matrices	FW Word fluency: COWAT, Letter/ Phonemic Fluency, Category/ Semantic Fluency, WLG
BD vs HC	Signif. differences	0%	33%	9%	23%	30%	40%	11%	8%
	Number of results	2	3	23	13	10	10	9	11
	ES of signif. Results	–	1 L	1 S, 1 L	3 L	3 L	1 M, 3 L	1 L	1 M
	Value/range of ES	–	1.04	0.13–1.04	0.93–1.40	0.83–1.16	0.68–1.74	0.87	0.76
NBD vs HC	Signif. differences	0%	100%	38%	30%	29%	50%	0%	25%
	Number of results	1	1	16	10	7	4	3	4
	ES of signif. Results	–	1 L	6 L	3 L	2 L	2 L	–	1 L
	Value/range of ES	–	1.05	0.99–2.87	1.52–2.38	1.19–1.25	1.24–1.27	–	1.66
NBD vs BD	Signif. differences	0%	0%	11%	20%	0%	50%	0%	0%
	Number of results	1	1	9	5	4	4	1	2
	ES of signif. Results	–	–	1 L	1 L	–	2 M	–	–
	Value/range of ES	–	–	1.47	1.01	–	0.61–0.64	–	–
NBD vs MS	Signif. differences	–	0%	13%	17%	0%	0%	0%	33%
	Number of results	0	1	8	12	6	2	5	3
	ES of signif. Results	–	–	1 L	2 L	–	–	–	1 L
	Value/range of ES	–	–	1.26	1.68–2.95	–	–	–	1.23

Abbreviations: *L* large, *M* medium, *S* small, *ES* effect size, *IQ* Intelligence Quotient, *MMSE* Mini Mental State Examination, *WMS MQ* Wechsler Memory Scales Memory Quotient. All other abbreviations as per Table 3

[a] the effect sizes of from significant results from studies with inadequate adjustment for multiple comparisons have only been calculated if the p-values remain significant after reviewer applied Bonferroni correction. Where p values were not significant after Bonferroni correction the results have been categorised as non-significant

Table 5 Case Series/Case Studies included in Systematic review

First Author, Year	Country	Cases	BD/NBD	Age	Gender	Disease duration (years)	BD criteria used	BD medication status	Longitudinal follow-up data	Psychological assessment measures
Segbedji, 2017	Morocco	1	NBD	39	nr	unclear	nr	Corticosteriods (post assessment)	nr; patient died 1 week post initial assessment	nr
Fisher et al., 2016	Australia	1	NBD	40	M	4	nr	Prednisolone, sodium valproate, colchicine and mycophenolate	Two brief prior assessments described, actual test data results not reported	Self-report mood scale (1–10)
Gabrielyan, 2015	Armenia	1	NBD	43	M	5	nr	Steroid treatment	Yes 3 month MMSE review following increase in steroid treatment	nr
Tosto, 2013	Italy	1	NBD	65	M	10 [a]	nr	Sertraline	Yes, re-assessment 3 years post initial presentation (MMSE and NPI)	NPI
Mimura, 2009	Japan	1	NBD	48	M	8 [a] 2	ISGBD	Not reported	nr	nr
Melillo, 2007	United Kingdom	1	NBD	36	F	18 [a]	nr	6-methylpredniso-lone, low-dose cyclo-phosphamide, 2-mer-captoethane sulpho-nate, prednisolone	Clinical follow up (including neurological assessment and brain imaging) but no cognitive re-assessment reported	nr
Park, 2007	Korea	1	NBD	50	M	<1	nr	Prednisolone (60 mg/day, tapered to 20 mg/day)	Yes, re-assessed 5 month post prednisolone therapy.	nr
Tsai, 2001	Taiwan	1	BD	30	M	5	nr	Prednisolone (5 years post onset)	Yes, re-assessed ~ 2 months post steroid treatment	nr
Hirohata et al., 1998; Suda, 1999	Japan	6	6 NBD	43–67	2 F 4 M	nr	ISGBD	All received methotrexate, various other including enoxaparus, bredy-nin, colchicine meth-ocarbam, reported	Yes for all, 6 and 12 month follow-up IQ assessment for all	nr
Oktem-Tanor, 1999	Turkey	12	12 NBD	20–33	12 M	nr	ISGBD	nr	Yes for all, between one and five follow-up assessments, at 6 months - 6 years, 8 months	nr
Zuliani, 1998	Italy	1	NBD	34	F	12	nr	Initially Clorambucyl then cortisone	Yes, re-assessed ~ 1 year post initial assessment	HAM-D, ZUNG
Kawanishi et al., 1995	Japan	1	NBD	50	M	4	nr	Prednisolone, antipsychotic drugs	Clinical follow-up over 7 month period, but no cognitive reassessment reported	nr
Arai et al., 1994	Japan	1	NBD	50	F	6	nr	Prednisolone	Clinical follow-up over ~ 6 year period but no cognitive reassessment reported	nr
Soyka, 1987	Germany	1	BD	39	F	8	nr	Amitriptylin, Methyl-prednisolone, Methy-pred, neuroleptic drugs	Yes, re-assessed after 4 years.	nr
	Italy	1	BD	58	M	31		Pimozide, ACTH	nr	nr

Table 5 (continued)

First Author, Year	Country	Cases	BD/NBD	Age	Gender	Disease duration (years)	BD criteria used	BD medication status	Longitudinal follow-up data	Psychological assessment measures
Bussone et al., 1982							Lehn-er's Cx.			
Yamada et al., 1978	Japan	6	6 NBD	24–44	2 F 4 M	5–21	nr	nr	Clinical follow-up over 1 to 5 years, but no cognitive reassessment reported for 5 cases, for 1 case the follow-up is unclear	nr
Wada, 1969	Japan	9	9 BD	28–39	3 F 6 M	2–10	nr	nr	nr	nr

Abbreviations: *ACTH* adrenocorticotropin, *HAM-D* Hamilton Depression Scale, *ISGBD* International Study Group for Behçet's Disease, *Lehner's Cx* Lehner's Criteria 1979, *nr* not reported, *NPI* Narcissistic Personality Inventory, *ZUNG* Zung Self Rating Depression Scale, *IQ* Intelligence Quotient. All other abbreviations as per Table 3

[a] since onset of gradual deterioration of cognitive function/symptom onset, not from diagnosis

memory (38%), verbal memory (33%), and executive functioning (27%). Rates of impairment on cognitive measures, "2 SD below 50th percentile in at least two tests" were provided. Based on this criterion, 31.4% of participants were categorised as having cognitive impairment. The rates of depression were reported as 60% based on BDI scores, with levels reported to be higher in participants with cognitive impairment.

Case Studies/Case Series

Study Characteristics, Quality and Bias Ratings

Details of the 17 included case study/series (reported across 18 papers) are shown in Table 5, including study details, demographics, and disease characteristics of the 46 participants. The majority of cases (34/46) were classified as having NBD. Information about the diagnostic criteria was only reported in four of the papers, however a further six papers provided information about the clinical symptoms and investigations that were used to make the diagnosis (Yamada, Kashiwamura, Nakamura, Ota, & Nakamura, 1978; Arai et al., 1994; Kawanishi et al., 1995; Soyka, 1987; Feng, Kuo, Chen, & Yeh, 2014; Fisher et al., 2016). The age of participants in the reports spanned 20 to 67 years and the majority of reported cases were male (M 34:12 F). Disease duration, or time since symptom onset, was not always reported, yet where available, this ranged from <1 to 21 years. Most papers provided information about the medication status of the participants. A range of medications were listed, with prednisolone the most commonly reported. The majority of reports described clinical follow-up over several years, with eight providing neurocognitive data from review assessments with participants at a later time point (often after medication treatment). Only three cases reported on psychological assessment in addition to neurocognitive assessment, and only two of those used standardised assessment measures.

The results of the bias and quality ratings are shown in Table 6. Based on these ratings, the majority of studies had problems with bias and quality. Information about blinding of cognitive assessors was not provided in the majority of studies. One study was rated as high for incomplete data reporting, while in a further eight this was unclear. Seven studies were rated as high for selective reporting bias, while a further five where rated as unclear. Almost all studies were rated as high for other forms of bias, and explanations for these ratings have been provided in the comments section in the table. This included unclear reporting of results, qualitative descriptions of the participants' mood without formal assessment, and a failure to address potential practice effects when reassessment was conducted over short time intervals. Appropriate use of, or reference to, normative data for the comparison of test results could not be found in four studies. Similarly,

Table 6 Risk of bias summary table for case series/studies

First Author, Year	Case Number	Blinding of Cognitive Assessor (Detection Bias)	Incomplete Data Reporting (Detection Bias)	Selective Reporting Rating (Reporting Bias)	Other Bias Rating	Appropriate use of/ reference to normative data for comparison	Premorbid ability level estimated/ accounted for	Comments
Segbedji, 2017	1	Low	Low	Low	Low	Yes, GCS standardised score	No	
Fisher et al., 2016	1	Unclear	Low	Unclear	High	Yes, for all tests except WMS-III Information & Orientation, L'Hermitte	Yes, Premorbid IQ estimate provided	Mood described but not formally assessed psychometrically.
Gabrielyan, 2015	1	Unclear	Low	Low	High	No	No	"affective disturbances" reported but mood not formally assessed .
Tosto, 2013	1	Unclear	Unclear	High	High	No	No	Reporting of results is unclear. Unclear when the testing for the reported results occurred.
Mimura, 2009	1	Unclear	Low	Low	High	Yes, for all tests except RAVLT – immediate recall	No, 'Average' used as the comparison point	Mood was not formally assessed despite description of "emotions appeared mildly flattened".
Melillo, 2007	1	Unclear	Unclear	Unclear	High	No	No	Severe depression reported but no description of psychometric tests being used to assess this.
Park, 2007	1	Unclear	Low	Unclear	High	Yes, for ROCF, HVLT, MMSE (reported %ile data)	No	Potential practice effects (5 months re-test interval) not considered or discussed.
Tsai, 2001	1	Unclear	Low	Low	High	No	No	Potential practice effects (2 months re-test interval) not considered or discussed. CASI is a dementia screen for elderly but is used here in a 30 year old.
Hirohata et al., 1998; Suda, 1999	1–6	Unclear	Low	Low	High	Yes, Wechsler standardised scoring but HDS-R norms provided with no interpretation	No	Potential practice effects (3 IQ tests in 12 months) not considered or discussed. IQ results reported on graphs with differing scales on the Y Axis may contribute to misleading interpretation. IQ data is reported but unclear which patient each set of results are attributable to.
Oktem-Tanor, 1999	1, 2, 5, 6, 7, 10, 11	Unclear	High	High	High	Yes, scored against age and education matched healthy control samples	No	Potential practice effects (6 to 12 month re-test intervals) not considered or discussed. Not all participants received all cognitive assessment measures at each assessment. Results not reported by test, but by overall domain functioning.
	3, 4, 8, 9, 12	Unclear	Unclear	High	High		No	

Table 6 (continued)

First Author, Year	Case Number	Blinding of Cognitive Assessor (Detection Bias)	Incomplete Data Reporting (Detection Bias)	Selective Reporting Rating (Reporting Bias)	Other Bias Rating	Appropriate use of/ reference to normative data for comparison	Premorbid ability level estimated/ accounted for	Comments
						Yes, scored against age and education matched healthy control samples		Inconsistencies in data reporting throughout paper.
Zuliani, 1998	1	Unclear	Unclear	High	High	Yes, Wechsler & Luria standardised scoring	No, but occupational information reported	A number of clinical tests conducted with casual reference to results - not specific data, e.g. "routine hematochemistry tests, substantially negative"
Kawanishi et al., 1995	1	Unclear	Unclear	High	High	Yes, Wechsler standardised scoring	No, but educational and occupational information reported	Most of psychiatric and behaviour symptoms are descriptive, does not appear to have been formal evaluation of these
Arai et al., 1994	1	Unclear	Unclear	High	High	Yes, Wechsler standardised scoring	No	Some of the case information provided is clinical opinion/observational.
Soyka, 1987		Unclear	Low	Unclear	High	Yes - Hawie standardised normative data	No, comment on pre-morbid abilities but not taken into account for interpretation of results.	Repeated reference to depression and also psychosis with no reports of assessing this formally, using psychometric tools.
Bussone et al., 1982	1	Unclear	Low	Low	Low	Yes - Wechsler standardised scoring	No	
Yamada et al., 1978	1, 3–6 2	Unclear Unclear	Unclear Unclear	High Unclear	High High	Unclear - no details Yes for Wechsler tests, unclear for other tests	Unclear - no details Unclear - no details	A large amount of the case information provided is clinical opinion/observational. It also does not appear that standardised assessments/ratings have been applied consistently across cases.
Wada, 1969	1 (2–9 cognitive data nr)	Unclear	Low	Low	Unclear	Yes - Wechsler standardised scoring	No	Results written in narrative, therefore determining the full extent of bias is difficult.

Abbreviations: *CASI* Cognitive Abilities Screening Instrument, *GCS* Glasgow Coma Score, *HDS-R* Revised Hasegawa's Dementia Scale, *IQ* Intelligence Quotient, *nr* not reported, *WMS – III* Wechsler Memory Scales Third Edition. All other abbreviations as per Table 3

Table 7 Neurocognitive assessment results of case series/study papers[a]

First Author, Year	Case	BD/ NBD	Normal, unimpaired or consistent with premorbid functioning	Abnormal, impaired or below premorbid functioning	Measures described but results not reported	Major neurocognitive findings as reported by study authors
Segbedji, 2017	1	NBD	nr	GCS	None	Sudden onset of impaired mental status
Fisher et al., 2016	1	NBD	WIAT-II – Word reading, TMT-A	WAIS-IV (Similarities, Information, BD, SS), Verbal Fluency (letters 'F', 'A', 'S', and animal naming), TMT B, WMS-IV (LMI & LM II), HVLT (learning, delayed recall, delayed DI), RCF (copy, 3-min & 30 min recall), Porteus Mazes – Vineland Revision.	None. Although two previous brief assessments described	"…majority of his performances fell in the Borderline to Extremely Low range of ability, suggestive of mild-to-moderate reductions in a number of areas. Difficulties were most apparent in the areas of working memory, concentration, processing speed, sequencing, letter verbal fluency (orthographical lexical retrieval), encoding on memory testing, and executive functioning (verbal abstract reasoning, planning, organization, problem solving, and strategy generation)."
Gabrielyan, 2015	1	NBD	nr	MMSE	None	"mild cognitive impairment"
Tosto, 2013	1	NBD	nr	MMSE	RAVLT, RCFT, TMT, Stroop Color-Word, Colored Progressive Matrices, FAB, Verbal fluency (letters 'F', 'A', 'S') & Animal Naming	"…subsequent neuropsychological assessments highlighted the gradual deterioration of his cognitive status... his behavioural disturbance worsened: depression and marked loss of interest in addition to his increasing anxiety and irritability".
Mimura, 2009	1	NBD	WAIS-R VIQ, PIQ, RCPM, RCFT Copy, WAB AQ	WMS-R Delayed recall, Verbal, General, RAVLT – delayed recall, RCFT – delayed recall, TMT B, WCST – categories 1 & perseveration, Word Fluency	Autobiographical Memory Interview	"…chronic and stable state of combination of amnesia and frontal/executive dysfunction. We concluded that the patient's clinical features were consistent with those of non-alcoholic Korsakoff syndrome resulted from neuro-Behcet's disease."
Melillo, 2007	1	NBD	nr	Luria, MMSE	Unclear	"… she had apraxia, agraphaesthesia and astereognosia" - no indication of formal tests used to assess this. Long-term memory loss also referred to in the case description. No overall conclusions about neurocognitive results.
Park, 2007	1	NBD	Digit span and letter cancellation, 'normal' (tests and exact results unclear). ROCFT – copy	MMSE, HVLT learning, delayed recall, ROCF delayed recall, Korean Famous Events Test	Digit span and letter cancellation (tests and exact results unclear).	"… selective anterograde amnesia with learning difficulty on the neuropsychological test without focal neurological deficits…. retrograde amnesia was minimal…. a very slight impairment in the know-ledge of events over the preceding months and years."
Tsai, 2001	1	BD	Alert to stimuli and spatial perception was preserved (tests not specified).	MMSE, CASI	None	The mental function … declined gradually over the past five years before the acute episode of midbrain lesion..... with the memory function being the most commonly and most severely affected." "However the CASI … did not show as specific or selective involvement on the frontal lobes."
Hirohata et al., 1998; Suda, 1999	1	NBD	HDS-R	HDS-R for Case 5. IQ measures cannot be attributed to individual	Orientation, memory and language functioning described but not tests reported	"memory disorder, mild articulation disorder"
	2	NBD	HDS-R			"disorientation", "low computing power" and "memory disorder / dysarthria" described

Table 7 (continued)

First Author, Year	Case	BD/ NBD	Normal, unimpaired or consistent with premorbid functioning	Abnormal, impaired or below premorbid functioning	Measures described but results not reported	Major neurocognitive findings as reported by study authors
	3	NBD	HDS-R	cases. Group averages		"disorientation" reported
	4	NBD	HDS-R	obtained from extracting		"memory disorder" reported
	5	NBD	nr	data from Fig. 2 indicate		"memory disorder" reported
	6	NBD	HDS-R	mildly to moderately reduced scores compared to test mean for VIQ, PIQ and Total IQ (Suda, 1999, Fig. 4).		"aphasia" described
Oktem-Tanor, 1999	1	NBD	Attention	Memory, Frontal functions	Orientation/inform-ation, language, visuospatialfunctions, directed spatial attention, reasoning, abstraction and arithmetic assessed but not included in results as authors state "these areas of cognition were almost always intact"	"Analysis of our data revealed a special pattern of cognitive decline with impaired memory, attention, and "frontal lobe functions" with poor motivation and personality change, in contrast to the relatively preserved linguistic, arithmetic, visuospatial, abstraction, and problem-solving abilities. The major finding in our series was memory impairment, not only because of its presence in all 12 patients, but also because it assumed first place among cognitive defects in terms of severity rating from moderate to severe disturbance."
	2	NBD	nr	Attention, Memory		
	3	NBD	nr	Attention, Memory, Frontal functions		
	4	NBD	nr	Attention, Memory, Frontal functions		
	5	NBD	Attention, Frontal functions	Memory		
	6	NBD	Attention	Memory		
	7	NBD	Attention	Memory		
	8	NBD	nr	Attention, Memory, Frontal functions		
	9	NBD	nr	Attention, Memory, Frontal functions		
	10	NBD	Attention	Memory		
	11	NBD	Attention	Memory		
	12	NBD	nr	Attention, Memory, Frontal functions		
Zuliani, 1998	1	NBD	IQ	nr	Luria Nebraska Neuropsychological Battery	"proper verbalization… a slight compromise of reasoning ability through symbols, abstraction and categorization, fixation memory (verbal and nonverbal) and attention." No details about which tests this was determined from, or specific scores
Kawanishi et al., 1995	1	NBD	Verbal IQ	Performance IQ	STM, orientation, aphasia, apraxia and agnosia skills referred to. Tests not reported.	"no evidence of an aphasia, apraxia, or agnosia… short-term memory disturbance characterized by disorientation to time… The present case illustrates the slowly progressive dementia, personality changes of a frontal lobe syndrome…"
Arai et al., 1994	1	NBD	nr	Verbal IQ	Short and long term memory, orientation, calculation, and inhibition referred to. Tests not reported.	"On admission she was demented (verbal IQ 78 by WAIS), with marked intellectual impairments, short- and long-term memory disturbances, disorientation, dyscalculia, affective incontinence and euphoria." "dementia had progressed gradually", "lacking inhibition".
Soyka, 1987	1	BD	nr	IQ	Language and memory functioning described. Tests not reported.	"poor/ reduced language/speech" as well as "disturbances/breaks in long-term memory and the ability to remember/memory", "clear indications of a

Table 7 (continued)

First Author, Year	Case	BD/ NBD	Normal, unimpaired or consistent with premorbid functioning	Abnormal, impaired or below premorbid functioning	Measures described but results not reported	Major neurocognitive findings as reported by study authors
						secondary organic performance reduction/degradation of performance".
Bussone et al., 1982	1	BD	IQ	nr	None	"His IQ was in the normal range".
Yamada et al., 1978	1	NBD	Not reported	"Memory impairment"	Unclear	"In the measurement of 1Q by WAIS, scores on the performance test were low in many cases partly because of psychiatric symptoms being present, but in the verbal test, scores were higher than expected and the calculating capacity was retained relatively well in many cases. This condition lasted from six months to four years."
	2	NBD	WAIS Verbal Test	WAIS Performance Test	Unclear	"Recent memory impairment was observed". Performance IQ on WAIS impaired. "An intelligence test revealed that his understanding and calculating capacity were almost normal. Aphasia and apraxia were not observed"
	3	NBD	nr	nr	Unclear	"Capacity for calculating was almost normal", "A decline in memory was noted"
	4	NBD	nr	nr	Unclear	Disorientation for time and space was observed, memory impairment, orientation for people retained.
	5	NBD	nr	nr	Unclear	Memory impairment
	6	NBD	nr	nr	Unclear	Memory impairment
Wada, 1969	1	BD	nr	WAIS	Orientation described but text not reported	"disorientation" reported
	2–9	BD	nr	nr	nr	nr

Abbreviations: *BD* Block Design, *CASI* Cognitive Abilities Screening Instrument, *DI* Discrimination Index, *FAB* Frontal Assessment Battery, *HDS-R* Revised Hasegawa's Dementia Scale, *LM-I* Logical Memory I, *LM-II* Logical Memory II (Recall), *nr* not reported, *MMSE* Mini Mental State Examination, *PIQ* Performance Intelligence Quotient, *STM* short term memory, *SS* Symbol Search, *VIQ* Verbal Intelligence Quotient, *WAB AQ* Western Aphasia Battery Aphasia Quotient, *WIAT-II* Wechsler Individual Achievement Test – Second Edition, *WMS – IV* Wechsler Memory Scales Fourth Edition. All other abbreviations as per Table 3

[a] for studies in which more than one set of neurocognitive data were presented (i.e. papers with review assessments), only the first or primary assessment results are presented here

participants premorbid level of functioning was not assessed or considered during interpretation in the majority of studies.

Neurocognitive Assessment Data

Table 7 provides a summary of the neurocognitive assessment results from the studies as well as the major findings reported by study authors. For papers reporting on more than one set of neurocognitive data for participants (i.e., those who conducted reviews over time), only the first set of reported data (i.e., assessment Time 1) was included in Table 7. The studies varied in the size and nature of the neurocognitive assessment measures applied. In three studies, just one basic measure or cognitive screening assessment result was provided, in seven papers only IQ assessment measure results were reported, while in four papers larger-scale neurocognitive batteries assessing multiple domains were reported.

When reported, all MMSE results were considered to be abnormal, impaired or below premorbid functioning on initial testing (seven assessment results). This contrasted with the HDS-R results (a similar measure to the MMSE) where five out of six participants were considered to have scored in the 'normal' range. The CASI, a cross-cultural dementia assessment tool, was utilised in one case study paper and was considered to indicate impaired functioning. Where IQ measures were reported the findings varied. In four single case papers, all reported IQ measures (total, index, scale or selected subtest scores) were considered to be abnormal, impaired or below premorbid functioning. In two single case papers the full-scale/total IQ measure was considered to be normal, unimpaired or consistent with premorbid functioning. In two single case papers, split index scales were reported with the Verbal IQ reported to be normal, unimpaired or consistent with premorbid functioning, and the Performance IQ considered abnormal, impaired or below premorbid functioning. While in the remaining six case series paper, half of the participants exhibited total IQ scores in the Average range of functioning, and half exhibited scores below the Average range. This ratio was also observed for both the Verbal and Non-verbal index scales in this paper (Hirohata, Suda, & Hashimoto, 1998; Suda, 1999). In the studies with broader testing batteries, neurocognitive functioning on assessment measures pertaining to memory functioning were most the frequently considered abnormal, impaired or below premorbid functioning. This was followed by measures of attention and "frontal" (executive) functioning and verbal fluency. Overall, of the cases with clearly documented formal neurocognitive test results, all participants except one (Bussone, La Mantia, Boiardi, & Giovannini, 1982) were described as have some degree of attenuated neurocognitive functioning.

Discussion

To our knowledge, this is the first attempt to systematically review the literature on neurocognitive functioning in BD and NBD. The review indicates that the majority of research in this area has examined neurocognitive outcomes through group comparisons (BD, NBD, MS and HC) and case study/series papers, with few group characterisation studies. The review highlights concerns about the methodology and quality of reporting in the published studies. Due to these limitations, drawing clear conclusions about neurocognitive functioning in BD and NBD was difficult, based on the currently available evidence, and prevented further synthesis of the data.

Results from group comparison studies suggest that neurocognitive changes in BD/NBD may be on a continuum, e.g., HC > BD > NBD. Neurocognitive attenuation in BD compared to HC is less robust, and not as frequently observed, compared to neurocognitive attenuation in NBD compared to HC. In both patient groups (BD and NBD), neurocognitive deficits, when observed, were more common in the domains of Gv Visual-spatial ability, Gsm Working Memory and Gc Acquired knowledge/crystallised ability, while additional deficits in Gs Processing Speed and Glr Long-term Memory Encoding and Retrieval where also observed in the NBD group. Our decision to use a theory-driven model to allocate neurological test results to specified domains (CHC model applied to clinical populations as per Jewsbury and colleagues) may impact on the comparability of the results to others in the field, particularly due to the subsuming of 'executive' test results into other domain constructs. However, it is noted that even when an executive functioning category or categories are demarcated, there is no clear consensus on which neurocognitive skills or tests fall under this sphere, and thus issues of comparability are a problem across the field and are not isolated to this review (Banich, 2009; Jewsbury et al., 2016; Goodall et al., 2018).

The available evidence suggests that the neurocognitive profiles of NBD and MS patients are largely comparable. Word fluency was the only cognitive construct where more than 25% of results suggested NBD performed significantly more poorly than MS. Both MS and NBD are disorders associated with multiple cerebral lesions that over time may develop, resolve, or change in size, and often occur in similar brain regions (Fisher et al., 2016; Gündüz et al., 2012; Koçer et al., 1999; Polman et al., 2012; Siva & Saip, 2009). While the underlying pathological causes of the lesions may be different, the similarity of the lesion locations may contribute to a similar profile of neurocognitive deficits being observed.

A large number of case studies/series explored neurocognitive functioning in the NBD patients, with fewer published case studies in BD. Like their group comparison counterparts, the literature in this area is impacted by a number of problems with study quality and reporting. Of most concern

is the high number of case studies/series papers that failed to adequately document which criteria were used to diagnose BD and NBD in their patients. In addition, few studies provided information regarding whether the neurocognitive assessor was aware of the diagnostic status of the patient. However, it is important to note that blinding to patient diagnosis is less common, and often not possible, in clinical settings. The majority of reported case studies did not assess for, or take into account, an estimate of the patient's premorbid level of functioning. This increases the likelihood of the over diagnosis of deficits, as scores falling below the Average range of functioning are likely considered as evidence of attenuated cognitive functioning, when, in fact, they may be consistent with the patient's 'premorbid' ability level.

In the vast majority of reported BD and NBD case studies some degree of attenuated neurocognitive functioning was described by the study authors. This may reflect the tendency of clinicians and study groups to submit cases for publication that are judged to be clinically interesting or noteworthy, rather than cases they see that are less interesting and more reflective of 'normal' functioning. However, it may also indicate that when patients are referred for neurocognitive assessment it is because there are clear functional concerns about their cognition, which are then confirmed through formal testing procedures. As a group, the case studies/series indicate that performance on intellectual testing measures may be reduced from the average range of functioning in some patients with BD and NBD. Attenuated functioning in specific neurocognitive domains has been most commonly observed in NBD case studies in the areas of memory, attention, verbal fluency and "frontal" or what is commonly referred to as executive functioning. Overall, this appears to parallel the results of the group comparison studies, as the majority of memory tasks are subsumed under Glr Long-term memory encoding and retrieval construct, attention under Gsm Working memory, and most executive functioning tasks into the Gs Processing Speed (switching and inhibition tasks) or Glr Long-term memory encoding and retrieval (updating) constructs.

There are likely to be multi-factorial causes of attenuated neurocognitive functioning in BD and NBD.The available evidence indicates that both BD and NBD sufferers exhibit higher rates of depression and anxiety than HC participants and that rates of depression and anxiety in BD and NBD sufferers are similar, with these psychological disorders known to impact on neurocognition (Allott, Fisher, Amminger, Goodall & Hetrick, 2016; Rock, Roiser, Riedel, & Blackwell, 2014; Castaneda, Tuulio-Henriksson, Marttunen, Suvisaari, & Lönnqvist, 2008). Both BD and NBD patients are also treated with medication, including corticosteroids, that may impact on neurocognitive functioning (Lupien, Gillin, & Hauger, 1999; Brown & Chandler, 2001) and may also increase the rates of psychological disorders (Patten, 2000). Cogently, the existence of cerebral lesions in NBD forms a plausible reason why neurocognitive dysfunction may be present in this patient group. This may also be the reason why neurocognitive attenuation is found at a higher prevalence in studies comparing NBD cohorts to controls, than studies of BD cohorts and controls. However, it is important to note that not all studies with BD groups included in this review outlined whether/how NBD was excluded in their BD cohort. Thus, the possibility cannot be ruled out that a proportion of BD participants in these studies have undiagnosed neurological manifestations, sometimes referred to as 'silent' NBD."

Further research into neurocognitive functioning in BD and NBD is recommended and the field would benefit from several larger sample studies with broad neurocognitive test batteries that are well designed, comparing functioning in BD and NBD to HC cohorts. Several independently conducted multi-centre studies, using similar (or identical) large testing batteries based on theory driven models of cognitive functioning, such as the CHC model, would strengthen the literature base and allow for more statistically accurate synthesis of the literature via methods such as meta-analysis. Efforts should be taken to minimise the key areas of bias identified in this review, fully report all results, and use neuropsychological measures that appropriately translate across language and cultural differences. Longitudinal larger sample group cohort studies would also provide information about the impact of BD/NBD on neurocognitive functioning over time, as presently information regarding the longitudinal neurocognitive impact of the disorder is provided only in the case study/series literature. Measurement and analysis of co-varying factors, such as psychological functioning and medication status (particularly corticosteroids) on cognition in BD and NBD would also be useful.

Acknowledgements We would like to thank Tony Monteleone, Michael Johnston, Marie-Christina Rieder and Roy Kessels for their translation assistance.

Compliance with Ethical Standards

Conflict of Interest Author A is also an author on one of the papers included in the systematic review.

References

Al-Dalaan, A. N., al Balaa, S. R., el Ramahi, K., Al-Kawi, Z., Bohlega, S., Bahabri, S., & al Janadi, M. A. (1994). Behçet's disease in Saudi Arabia. *Journal of Rhematology, 21*(4), 658–661.

Alibaz-Oner, F., Sawalha, A. H., & Direskeneli, H. (2018). Management of Behçet's disease. *Current Opinion in Rheumatology, 30*(3), 238–242.

Allott, K., Fisher, C. A., Amminger, G. P., Goodall, J., & Hetrick, S. (2016). Characterizing neurocognitive impairment in young people

with major depression: State, trait, or scar? *Brain and Behaviour, 6*(10), e00527. https://doi.org/10.1002/brb3.527

Arai, T., Mizukami, K., Sasaki, M., Tanaka, Y., Shiraishi, H., Horiguchi, H., et al. (1994). Clinicopathological study on a case of neuro-Behçet's disease: In special reference to MRI, SPECT and neuropathological findings. *The Japanese Journal of Psychiatry and Neurology, 48*(1), 77–84.

Bang, D., Lee, J. H., Lee, E. S., Lee, S., Choi, J. S., Kim, Y. K., ... & Park, S. D. (2001). Epidemiologic and clinical survey of Behçet's disease in Korea: The first multicenter study. *Journal of Korean Medical Science, 16*(5), 615–618.

Banich, M. T. (2009). Executive function. The search for an integrated account. *Current Directions in Psychological Science, 18*(2), 89–94. https://doi.org/10.1111/j.1467-8721.2009.01615.x

Beeldman, E., Raaphorst, J., Twennaar, M. K., De Visser, M., Schmand, B. A., & De Haan, R. J. (2016). The cognitive profile of ALS: A systematic review and meta-analysis update. *Journal of Neurology, Neurosurgery and Psychiatry, 87*(6), 611–619.

Bingol, A., Topcular, B., Yildiz, S., Kaya, P., Tutuncu, M., Demirci, O., ... & Siva, A. (2011). Is Cognitive Dysfunction in Multiple Sclerosis and Neuro-Behçet's Disease Similar? Paper presented at the Twenty-first Meeting of the European Neurological Society 28–31 May, Lisbon, Portugal. *Journal of Neurology, 258*(Supplement 1), S102. https://doi.org/10.1007/s00415-011-6026-9

Borson, S. (1982). Behçet's disease as psychiatric disorder: A case report. *American Journal of Psychiatry, 139*(10), 1348–1349.

Brown, E. S., & Chandler, P. A. (2001). Mood and cognitive changes during systemic corticosteroid therapy. *Primary Care Companion to the Journal of Clinical Psychiatry, 3*(1), 17–21.

Bussone, G., La Mantia, L., Boiardi, A., & Giovannini, P. (1982). Chorea in Behçet's syndrome. *Journal of Neurology, 227*(2), 89–92.

Castaneda, A. E., Tuulio-Henriksson, A., Marttunen, M., Suvisaari, J., & Lönnqvist, J. (2008). A review on cognitive impairments in depressive and anxiety disorders with a focus on young adults. *Journal of Affective Disorders, 106*(1), 1–27.

Cavaco, S., Da Silva, A. M., Pinto, P., Coutinho, E., Santos, E., Bettencourt, A., ... & Carvalho, L. (2009). Cognitive functioning in Behçet's disease. *Annals of the New York Academy of Sciences, 1173*, 217–26.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Dalvi, S. R., Yildirim, R., & Yazici, Y. (2012). Behçet's Syndrome. *Drugs, 72*(17), 2223–2241.

Davatchi, F. (2012). Diagnosis/classification criteria for behcet's disease. *Pathology Research International, 2012*. https://doi.org/10.1155/2012/607921

Davatchi, F., Assaad-Khalil, S., Calamia, K. T., Crook, J. E., Sadeghi-Abdollahi, B., Schirmer, M., ... & Alekberova, Z. S. (2014). The international criteria for Behçet's disease (ICBD): A collaborative study of 27 countries on the sensitivity and specificity of the new criteria. *Journal of the European Academy of Dermatology and Venereology, 28*(3), 338–347.

Davatchi, F., Chams-Davatchi, C., Ghodsi, Z., Shahram, F., Nadji, A., Shams, H., ... & Sadeghi-Abdolahi, B. (2011). Diagnostic value of Pathergy test in Behçet's disease according to the change of incidence over the time. *Clinical Rheumatology, 30*(9), 1151–1155.

Davatchi, F., Shahram, F., Chams-davatchi, C., Shams, H., Nadji, A., Akhlaghi, M., ... & Abdollahi, B. S. (2010). Behçet's disease in Iran: Analysis of 6500 cases. *International Journal of Rheumatic Diseases, 13*, 367–373.

Domingos, J., Ferrão, C., Ramalho, J., Rodrigues, T., Moreira, B., Santos, E., ... & Vasconcelos, C. (2015). Characteristics of neuro-Behçet's disease in a case-series from a single Centre in Northern Portugal. *European Neurology, 73*(5–6), 321–328.

Dutra, L. A., de Souza, A. W. S., Alessi, H., Guedes, B. D. V. S., Braga-Neto, P., Pedroso, J. L., ... & Barsottini, O. G. P. (2013). Cognitive impairment in Brazilian patients with Behçet's disease occurs independently of neurologic manifestation. *Journal of the Neurological Sciences, 327*(1–2), 1–5.

Erberk-Ozen, N., Birol, A., Boratav, C., & Kocak, M. (2006). Executive dysfunctions and depression in Behçet's disease without explicit neurological involvement. *Psychiatry and Clinical Neurosciences, 60*, 465–472.

Erkol, G., Vural, M., Karantay, F., & Uluduz, D. (2009). Event-related potentials and paced auditory serial addition test (PASAT) evaluation in Behçet's syndrome with subclinical neural involvement. *Nobel Medicus, 5*(2), 29–34.

Evans, V. C., Iverson, G. L., Yatham, L. N., & Lam, R. W. (2014). The relationship between neurocognitive and psychosocial functioning in major depressive disorder: A systematic review. *Journal of Clinical Psychiatry, 75*(12), 1359–1370.

Farahangiz, S., Sarhadi, S., Safari, A., & Borhani-Haghighi, A. (2012). Magnetic resonance imaging findings and outcome of neuro-Behçet's disease: The predictive factors. *International Journal of Rheumatic Diseases, 15*(6), 142–149.

Feng, H., Kuo, S., Chen, C., & Yeh, Y. (2014). Nocturnal bruxism in a patient with Behçet disease and posttraumatic stress disorder successfully treated with gabapentin. *Clinical Neuropharmacology, 37*(6), 190–191.

Fisher, C. A., Sewell, K., & Baker, A. (2016). Chronic behavior disturbance and neurocognitive deficits in neuro-Behçet's disease: A case study. *Neurocase, 22*(3), 332–338.

Gabrielyan, I., Harutyunyan, K., Avagyan, G., Amirjanyan, H., Khachaturyan, S., Hambardzumyan, H., & Manvelyan, H. M. (2015). Neuro-Behcet's disease: A case report. *European Journal of Neurology, 22* (Suppl. 1, 510.

Ganzer, F., Bröning, S., Kraft, S., Sack, P. M., & Thomasius, R. (2016). Weighing the evidence: A systematic review on long-term neurocognitive effects of Cannabis use in abstinent adolescents and adults. *Neuropsychology Review, 26*(2), 186–222.

Gates, N. J., & March, E. G. (2016). A Neuropsychologist's guide to undertaking a systematic review for publication: Making the Most of PRISMA guidelines. *Neuropsychology Review, 26*(2), 109–120.

Goodall, J., Fisher, C., Hetrick, S., Phillips, L., Parrish, E. M., & Allott, K. (2018). Neurocognitive functioning in depressed young people: A systematic review and meta-analysis. *Neuropsychology Review, 28*(2), 216–213. https://doi.org/10.1007/s11065-018-9373-9

Gündüz, T., Emir, Ö., Kürtüncü, M., Mutlu, M., Tumaç, A., Akca, S., et al. (2012). Cognitive impairment in neuro-Behcet's disease and multiple sclerosis: A comparative study. *International Journal of Neuroscience, 122*(11), 650–656. https://doi.org/10.3109/00207454.2012.704454

Haghighi, A. B., Sarhadi, S., & Farahangiz, S. (2011). MRI findings of neuro-Behçet's disease. *Clinical Rheumatology, 30*, 765–770.

Hatemi, G., Yazici, Y., & Yazici, H. (2013). Behçet's Syndrome. *Rheumatic Disease Clinics of North America, 39*, 245–261.

Hernández, F. J. G., Medina, C. O., Romero, L. M., Román, J. S., Solís, D. G., Franco-Baux, J. R., Merchán, J. M. (2002). Utilidad de La SPECT Cerebral Con HMPAO-99mTc Y de Los Tests Psicométricos En El Diagnóstico de La Afección Neurológica de La Enfermedad de Behçet. *Medicina Clinica, 119*(12), 446–450.

Higgins, J. P. T., Altman, D. C., & Sterne, J. A. C. (2017). Chapter 8: Assessing risk of Bias in included studies. In M. S. Cumpston, J. P. T. Higgins, R. Churchill, & J. Chandler (Eds.), *Cochrane handbook for systematic reviews of interventions version 5.2.0 (sections 8:1–8:73). Cochrane* (p. 2017).

Hirohata, S., Suda, H., & Hashimoto, T. (1998). Low-dose weekly methotrexate for progressive neuropsychiatric manifestations in Behçet's disease. *Journal of the Neurological Sciences, 159*(2), 181–185.

Hoelzle, J.B. (2008). Neuropsychological assessment and the Cattell-horn-Carroll (CHC) cognitive abilities model. Theses and

Dissertations, 1192, University of Toledo, http://utdr.utoledo.edu/theses-dissertations/1192.

Ideguchi, H., Suda, A., Takeno, M., Kirino, Y., Ihata, A., Ueda, A., ... & Ishigatsubo, Y. (2010). Neurological manifestations of Behçet's disease in Japan: A study of 54 patients. *Journal of Neurology, 257*, 1012–1020, https://doi.org/10.1007/s00415-010-5454-2.

International Study Group for Behçet's Disease. (1990). Criteria for diagnosis of Behçet's disease. International study Group for Behçet's disease. *Lancet, 335*(8697), 1078–1080.

Jewsbury, P. A., & Bowden, S. C. (2017). Construct validity of fluency and implications for the factorial structure of memory. *Journal of Psychoeducational Assessment, 35*(5), 460–481.

Jewsbury, P. A., Bowden, S. C., & Duff, K. (2017). The Cattell–Horn–Carroll model of cognition for clinical assessment. *Journal of Psychoeducational Assessment, 35*(6), 547–567. https://doi.org/10.1177/0734282916651360

Jewsbury, P. A., Bowden, S. C., & Strauss, M. E. (2016). Integrating the switching, inhibition, and updating model of executive function with the Cattell-horn-Carroll model. *Journal of Experimental Psychology: General, 145*(2), 220–245.

Kale, N., Agaoglu, J., Icen, M., Yazici, I., & Tanik, O. (2008). The presentation of headache in neuro-Behçet's disease: A case-series. *Headache, 49*(3), 467–470.

Kalra, S., Silman, A., Akman-Demir, G., Bohlega, S., Borhani-Haghighi, A., Constantinescu, C., ... & Siva, A. (2014). Diagnosis and Management of Neuro-Behçet's disease: International consensus recommendations. *Journal of Neurology, 261*(9), 1662–1676.

Kawanishi, Y., Hori, T., Suzuki, T., Mizukami, K., Sasaki, M., & Shiraishi, H. (1995). A clinical case of neuro-Behçet syndrome without Mucocutaneo-ocular symptoms: Relationship with brain stem encephalitis. *Psychiatry and Clinical Neurosciences, 49*(5), 263–266.

Kidd, D. P. (2017). Neurological complications of Behçet's syndrome. *Journal of Neurology, 264*(10), 2178–2183.

Kikuchi, H., Sawada, T., Okada, M., Takeno, M., Kuwana, M., Ishigatsubo, Y., ... & Hirohata, S. (2017). Recommendations for the management of neuro-Behçet's disease by the Japanese national research committee for Behçet's disease. *Journal of the Neurological Sciences, 381*, 646–647.

Koçer, N., Islak, C., Siva, A., Saip, S., Akman, C., Kantarci, O., & Hamuryudan, V. (1999). CNS Involvement in Neuro-Behçet Syndrome: An MR Study. *American Journal of Neuroradiology, 20*(6), 1015–1024.

Kronborg, C., Mahar, P. D., & Kelly, R. (2014). Should we keep changing the diagnostic criteria for Behçet's disease? *Dermatology, 228*, 1–4.

Larrabee, G. J. (2014). Test validity and performance validity: Considerations in providing a framework for development of an ability-focused neuropsychological test battery. *Archives of Clinical Neuropsychology, 29*(7), 695–714.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., ... & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *British Medical Journal, 6*(7), e1000100, https://doi.org/10.1136/bmj.b2700

Lupien, S. J., Gillin, C. J., & Hauger, R. L. (1999). Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: A dose–response study in humans. *Behavioral Neuroscience, 113*(3), 420–430.

Madanat, W., Fayyad, F., Verity, D., & Zureikat, H. (2000). Influence of sex on Behçet's disease in Jordan. In D. Bang, E. Lee, & S. Lee (Eds.), *Behçet's disease* (pp. 90–93). Seoul: Design Mecca Publishing.

Mason, R. M., & Barnes, C. G. (1969). Behçet's syndrome with arthritis. *Annals of the Rheumatic Diseases, 28*(2), 95.

Matsui, T., Ishida, T., Tono, T., Yoshida, T., Sato, S., & Hirohata, S. (2010). An Attache of cute neuro-Behçet's disease during the course of chronic progressive neuro-Behçet's disease: Report of two cases. *Modern Rheumatology, 20*, 621–626.

Melillo, N., Sangle, S., Stanford, M. R., Andrews, T. C., & D'Cruz, D. P. (2007). Low-dose intra-venous cyclophosphamide therapy in a patient with neurological complications of Behçet's disease. *Clinical Rheumatology., 26*(8), 1365–1367. https://doi.org/10.1007/s10067-006-0385-0

Mimura, M., Kato, M., & Kashima, H. (2009). Neuro-Behcet's disease presenting with amnesia and frontal dysfunction. *Clinical Neurology and Neurosurgery, 111*(10), 889–892. https://doi.org/10.1016/j.clineuro.2009.08.002

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *British Medical Journal, 339*(b2535), 332–336.

Monastero, R., Camarda, C., Pipia, C., Lopez, G., Camarda, L. K., Baiamonte, V., ... & Camarda, R. (2004). Cognitive impairment in Behçet's disease patients without overt neurological involvement. *Journal of the Neurological Sciences, 220*, 99–104.

Noel, N., Bernard, R., Wechsler, B., Resche-Rigon, M., Depaz, R., Le Thi Huong Boutin, D., ... & Saadoun, D. (2014). Long term outcomes in neuro-Behçet's disease. *Arthritis & Rheumatology, 66*(5), 1306–14.

Nurova, B., Kurtuncu, M., Coban, A., Birday, E., & Eraksoy, M. (2014). Computer assisted cognitive rehabilitation in patients with multiple sclerosis and parenchymal neuro-Behçet's disease. *Journal of Neurology, 261*(supplement 1), S246.

O'Duffy, J. D., & Goldstein, N. P. (1976). Neurologic involvement in seven patients with Behçet's disease. *The American Journal of Medicine, 61*(2), 170–178.

Oktem-Tanör, O., Baykan-Kurt, B., Gürvit, I. H., Akman-Demir, G., & Serdaroğlu, P. (1999). Neuropsychological follow-up of 12 patients with neuro-Behçet disease. *Journal of Neurology, 246*, 113–119. https://doi.org/10.1007/s004150050317

Özdemir, D. F., Özsoylar, G., Candansayar, S., Coşar, B., & Önder, M. (2004). Psychiatric findings related to neurological complications in Behçet's disease: A short review and a case presentation. *International Journal of Psychiatry in Clinical Practice, 8*, 185–190.

Özen, N. E., Birol, A., Boratav, C., & Koçak, M. (2004). Nörolojik Tutulumu Olmayan Behçet Hastalarýnda Biliþsel Bozukluklar. *Klinik Psikiyatri, 7*, 187–198.

Ozguler, Y., & Hatemi, G. (2016). Management of Behçet's syndrome. *Current Opinion in Rheumatology, 28*(1), 45–50.

Patten, S. B. (2000). Exogenous corticosteroids and major depression in the general population. *Journal of Psychosomatic Research, 49*(6), 447–449.

Park, K. W., Seo, J. H., & Yoon, G. U. (2007). Selective anterograde amnesia with thalamus and hippocampal lesions in neuro-Behcet's disease. *Clinical Neurology and Neurosurgery, 109*(5), 470–473. https://doi.org/10.1016/j.clineuro.2007.02.016

Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., ... & Lublin, F. D. (2012). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of Neurology, 69*(2), 292–302.

Remmers, E. F., Cosan, F., Kirino, Y., Ombrello, M. J., Abaci, N., Satorius, C., ... & Aglar, O. (2010). Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nature Genetics, 42*(8), 698–702.

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine, 44*(10), 2029–2040.

Ségbédji, K. K. F., Abdulrazak, S., Nubukpo-Guménu, A. A., Benzagmout, M., Aggouri, M., Chakour, K., & Chaoui, M. F.

(2017). A fatal case of probable Neurobehcet. *Interdisciplinary Neurosurgery: Advanced Techniques and Case Management, 8*, 14–16. https://doi.org/10.1016/j.inat.2017.01.008

Shahram, F., Latifi, G., Noroozian, M., Ghadiri, N.K., Abdollahi, B.S., Mohammad, H.F., & Davatchi, F. (2010). 14th International Conference on Behçet's Disease, London, United Kingdom. In *Cognitive Impairment in Patients with Behçet's Disease: A Case-Control Study*, S159. London, United Kingdom.

Shakir, R. A., Sulaiman, K., Kahn, R. A., & Rudwan, M. (1990). Neurological presentation of neuro-Behçet's syndrome: clinical categories. *European Neurology*. https://doi.org/10.1159/000117356

Siva, A., & Saip, S. (2009). The Spectrum of nervous system involvement in Behçet's syndrome and its differential diagnosis. *Journal of Neurology, 256*(4), 513–529.

Soyka, M. (1987). Rezidivierende Verwirrtheitszustande Bei Zerebraler Manifestation Eines Morbus Behçet. *Nervenarzt, 58*, 388–391.

Sucullu Karadag, Y., Kurt, P., Sahin, K., Karaaslan, Y., Oztekin, N., & Ak, F. (2014). Cognitive impairment not only in NeuroBehçet but also for all Behçet disease phenotypes. *Journal of Neurological Sciences, 31*(3), 511–520.

Suda, H. (1999). Low-dose weekly methotrexate therapy for progressive neuro-Behçet's disease. *Japanese Journal of Clinical Immunology, 22*(1), 13–22.

Tosto, G., Letteri, F., Canevelli, M., & Bruno, G. (2013). Efficacy of sertraline in a patient with neuro-Beçhet's disease. Journal of Clinical Neuroscience, 20 (6). Elsevier Ltd: 896–897. https://doi.org/10.1016/j.jocn.2012.05.044.

Topicular, B., Bingol, A., Kaya, P., Yildiz, S., Uluduz, D., Siva, A., & Saip, S. (2011). Cognitive dysfunction in neuro-Behcet's disease. *Journal of Neurology, 258*(Suppl 1), S102.

Tsai, K. Y., Chen, J. T., Guo, W. Y., Fuh, J. L., & Liao, K. K. (2001). Behcet's disease with pathologic laughing and crying: A case report and literature review. *Acta Neurologica Taiwanica, 10*(4), 271–276.

Tursen, U., Gurler, A., & Boyvat, A. (2006). Evaluation of clinical findings according to sex in 2313 Turkish patients with Behçet's disease. *International Journal of Dermatology, 42*, 346–351.

Türsen, Ü., & Türsen, B. (2014). Treatment options in Behçet's disease. *Global Journal of Dermatology & Venereology, 2*, 27–49.

Uğuz, F., Dursun, R., Kaya, N., & Cilli, A. S. (2007). Quality of life in patients with Behçet's disease: The impact of major depression. *General Hospital Psychiatry, 29*, 21–24.

Valesini, G., Pezzi, P. P., Catarinelli, G., Accorinti, M., & Priori, R. (1991). Clinical manifestations of Behçet's disease in Italy: Study of 155 patients at Rome University. In J. D. O'Duffy & E. Kokmen (Eds.), *Behçet's disease basic and clinical aspects* (pp. 279–289). New York: Marcel Decker Inc..

Wada, T., & Takahashi, T. (1969). Psychiatric signs of Behcet's disease. *Nihon Rinsho (Japanese Journal of Clinical Medicine), 27*(9), 2328–2332.

Yamada, M., Kashiwamura, K., Nakamura, Y., Ota, T., & Nakamura, K. (1978). On psychiatric symptoms of neuro-Behçet's syndrome. *Folia Psychiatrica et Neurologica Japonica, 32*(2), 191–197.

Zayed, H., Effat, D., Nawito, Z., Abdou, A. A., Nasr El Din, M., El-Refaei, S., … Abu El-Einin, K. (2011). Silent central nervous system involvement in Egyptian Behçet's disease patients: clinical, psychiatric, and neuroimaging evaluation. *Clinical Rheumatology, 30*, 1173–1180.

Zuliani, C., Boni, P., & Fattorello, C. (1998). Valutazione Neuropsicologica in Un Caso Di Neuro-Behcet. *Rivista Medica, 4*(1), 67–70.